PATENT
CUSTOMER No. 32425

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
John HARLEY and
Judith JAMES

Serial No.: 08/781,296

Filed: January 13, 1997

For: DIAGNOSTICS AND THERAPY OF EPSTEIN-BARR VIRUS IN AUTOIMMUNE DISORDERS

Group Art Unit: 1631

Examiner: Lori A. Clow

Atty. Dkt. No.: OMRF:050US/SLH

ELECTRONIC FILING SUBMISSION

Date of Filing: June 30, 2006

BRIEF ON APPEAL

TABLE OF CONTENTS

PATENT
CUSTOMER No. 32425

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
John HARLEY and
Judith JAMES

Serial No.: 08/781,296

Filed: January 13, 1997

For: DIAGNOSTICS AND THERAPY OF EPSTEIN-BARR VIRUS IN AUTOIMMUNE DISORDERS

Group Art Unit: 1631

Examiner: Lori A. Clow

Atty. Dkt. No.: OMRF:050US

BRIEF ON APPEAL

MAIL STOP APPEAL BRIEF - PATENTS
Commissioner of Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This Brief on Appeal is filed in response to the final Office Action mailed on February 3, 2006, regarding the above-captioned application. A Notice of Appeal was filed on April 3, 2006, making this brief due on June 3, 2006. The fee for this brief is included herewith; if the fee is missing or deficient, appellants authorize the Commissioner to debit Fulbright & Jaworski L.L.P. Deposit Account No. 55-1212/OMRF:050US/SLH.

Please date stamp the enclosed postcard as evidence of receipt.

I. Real Party in Interest

The real party in interest of this application is the assignee, the Oklahoma Medical Research Foundation, Oklahoma City, OK.

II. Related Appeals and Interferences

There are no known related appeals or interferences (see Appendix C).

III. Status of Claims

Claims 1-24 were filed with the original application, and claims 25-40. Claims 1-26 and 28 have been canceled, and claims 30-40 are withdrawn from consideration due to a restriction requirement/election of species. Thus, claims 27 and 29 and are under examination, stand rejected, and are appealed (see Appendix A).

IV. Status of Amendments

No unentered amendments have been offered after the final rejection.

V. Summary of Invention

The present invention is directed to peptides of up to about 40 amino acids that retain an entire native epitope or portions sufficient to react with an autoantibody (Specification at page 20, lines 26-36) and comprising sequences from Table 5, at page 48 of the Specification.

VI. Issues

Whether claims 27 and 29 are properly rejected under 35 U.S.C. §112, first paragraph, for alleged lack of written description.

Whether claims 27 and 29 are properly rejected under 35 U.S.C. §102(b) over Chen *et al.* (Exhibit 1).

VII. Grouping of Claims

The claims stand or fall together.

VIII. Argument

A. Standard of Review

As an initial matter, appellant notes that findings of fact and conclusions of law by the U.S. Patent and Trademark Office must be made in accordance with the Administrative Procedure Act, 5 U.S.C. § 706(A), (E), 1994, and *Dickinson v. Zurko*, 527 U.S. 150, 158 (1999). Moreover, the Federal Circuit has held that findings of fact by the Board of Patent Appeals and Interferences must be supported by "substantial evidence" within the record. *In re Gartside*, 203 F.3d 1305, 1315 (Fed. Cir. 2000). In *In re Gartside*, the Federal Circuit stated that "the 'substantial evidence' standard asks whether a reasonable fact finder could have arrived at the agency's decision." *Id.* at 1312. Accordingly, it necessarily follows that an Examiner's position on Appeal must be supported by "substantial evidence" within the record in order to be upheld by the Board of Patent Appeals and Interferences.

B. Rejection Under 35 U.S.C. §112, First Paragraph

Claims 27 and 29 stand rejected under the first paragraph of §112 as lacking an adequate written description. The examiner has focused on two aspects of the claims – sequences in addition to those recited in the claims and combinations of peptides – in advancing the rejection. Appellants again traverse both aspects of the rejection.

Additional Sequences. Appellants submit that there is no basis in any written description case law, particularly in *Vas-Cath*, that would justify limiting the claims to only those particular sequences set forth in the application – in effect, a "consisting of" claim. Appellants

submit that the standard for written description is whether one of ordinary skill in the art would readily apprehend that applicants were in possession of the invention as claimed at the time of filing. A description as filed is presumed to be adequate, unless or until sufficient evidence or reasoning to the contrary has been presented by the examiner to rebut the presumption. See, *e.g.*, *In re Marzocchi*, 439 F.2d 220, 224, 169 USPQ 367, 370 (CCPA 1971). The examiner, therefore, must have a reasonable basis to challenge the adequacy of the written description. The examiner has the initial burden of presenting by a preponderance of evidence why a person skilled in the art would not recognize in an applicant's disclosure a description of the invention defined by the claims. *In re Wertheim*, 541 F.2d at 263, 191 USPQ at 97. See also MPEP §2163.04.

Again, as pointed out previously, the present application describes a variety of peptides with varying lengths, and those of skill in the art would readily understand that these recited sequences could be included within larger peptide segments (but not longer than about 40 amino acids as recited) while still accomplishing the goals of the present invention – binding to autoantibodies. The examiner has made no effort to establish why one of skill in the art would not immediately comprehend that sequences beyond those recited could be included, and that these sequences could be virtually ***any other peptide sequence*** imaginable. Rather, in first raising of this rejection in the next to last office action, the examiner provided a single case law cite that appellants easily distinguished in their last response. And in the final Office Action, the examiner continues to make no effort to tie the rule of law in the cited cases to the fact pattern here. Rather, the examiner simply provides quotations from cases that do nothing more than restate legal tenets – that is clearly insufficient to establish the propriety of a rejection ***that is based on an examination of the unique facts of each case.***

Indeed, the examiner here is adopting an overly broad application of prior decisions just because appellants' claims recite sequences. That practice has been quite clearly admonished by the Federal Circuit, however. The recent decision in *Capron v. Eshhar v. Dudas*, 418 F.3d 1349 (Fed. Cir. 2005) discusses that there is no *per se* requirement under written description that a specification contain a detailed discussion of genetic elements where those elements are well known to those in the field and/or not essential to defining the novel aspects of the invention:

> Both parties argue that the Board misconstrued precedent, and that precedent does not establish a *per se* rule requiring nucleotide-by-nucleotide re-analysis when the structure of the component DNA segments is already known, or readily determined by known procedures. The "written description" requirement implements the principle that a patent must describe the technology that is sought to be patented; the requirement serves both to satisfy the inventor's obligation to disclose the technologic knowledge upon which the patent is based, and to demonstrate that the patentee was in possession of the invention that is claimed. See *Enzo Biochem*, 296 F.3d at 1330 (the written description requirement "is the *quid pro quo* of the patent system; the public must receive meaningful disclosure in exchange for being excluded from practicing the invention for a limited period of time"); *Reiffin v. Microsoft Corp.*, 214 F.3d 1342, 1345-46 (Fed. Cir. 2000) (the purpose of the written description requirement "is to ensure that the scope of the right to exclude . . . does not overreach the scope of the inventor's contribution to the field of art as described in the patent specification"); *In re Barker*, 559 F.2d 588, 592 n.4 (CCPA 1977) (the goal of the written description requirement is "to clearly convey the information that an applicant has invented the subject matter which is claimed"). The written description requirement thus satisfies the policy premises of the law, whereby the inventor's technical/scientific advance is added to the body of knowledge, as consideration for the grant of patent exclusivity.
>
> The descriptive text needed to meet these requirements varies with the nature and scope of the invention at issue, and with the scientific and technologic knowledge already in existence. The law must be applied to each invention that enters the patent process, for each patented advance is novel in relation to the state of the science. Since the law is applied to each invention in view of the state of relevant knowledge, its application will vary with differences in the state of knowledge in the field and differences in the predictability of the science.
>
> For the chimeric genes of the Capon and Eshhar inventions, the law must take cognizance of the scientific facts. The Board erred in refusing to consider the state of the scientific knowledge, as explained by both parties, and in declining to consider the separate scope of each of the claims. None of the cases to which the Board attributes the requirement of total DNA re-analysis, *i.e.*, *Regents v. Lilly*, *Fiers v. Revel*, *Amgen*, or *Enzo Biochem*, require a re-description of what was already known. In *Lilly*, 119 F.3d at 1567, the cDNA for human insulin had never been characterized. Similarly in *Fiers*, 984

F.2d at 1171, much of the DNA sought to be claimed was of unknown structure, whereby this court viewed the breadth of the claims as embracing a "wish" or research "plan." In *Amgen*, 927 F.2d at 1206, the court explained that a novel gene was not adequately characterized by its biological function alone because such a description would represent a mere "wish to know the identity" of the novel material. In *Enzo Biochem*, 296 F.3d at 1326, this court reaffirmed that deposit of a physical sample may replace words when description is beyond present scientific capability. In *Amgen Inc. v. Hoechst Marion Roussel, Inc.*, 314 F.3d 1313, 1332 (Fed. Cir. 2003) the court explained further that the written description requirement may be satisfied "if in the knowledge of the art the disclosed function is sufficiently correlated to a particular, known structure." These evolving principles were applied in *Noelle v. Lederman*, 355 F.3d 1343, 1349 (Fed. Cir. 2004), where the court affirmed that the human antibody there at issue was not adequately described by the structure and function of the mouse antigen; and in *University of Rochester v. G.D. Searle & Co.*, 358 F.3d 916, 925-26 (Fed. Cir. 2004), where the court affirmed that the description of the COX-2 enzyme did not serve to describe unknown compounds capable of selectively inhibiting the enzyme.

The "written description" requirement must be applied in the context of the particular invention and the state of the knowledge. The Board's rule that the nucleotide sequences of the chimeric genes must be fully presented, although the nucleotide sequences of the component DNA are known, is an inappropriate generalization. When the prior art includes the nucleotide information, precedent does not set a *per se* rule that the information must be determined afresh. Both parties state that a person experienced in the field of this invention would know that these known DNA segments would retain their DNA sequences when linked by known methods. Both parties explain that their invention is not in discovering which DNA segments are related to the immune response, for that is in the prior art, but in the novel combination of the DNA segments to achieve a novel result.

The "written description" requirement states that the patentee must describe the invention; it does not state that every invention must be described in the same way. As each field evolves, the balance also evolves between what is known and what is added by each inventive contribution. Both Eshhar and Capon explain that this invention does not concern the discovery of gene function or structure, as in Lilly. The chimeric genes here at issue are prepared from known DNA sequences of known function. The Board's requirement that these sequences must be analyzed and reported in the specification does not add descriptive substance. The Board erred in holding that the specifications do not meet the written description requirement because they do not reiterate the structure or formula or chemical name for the nucleotide sequences of the claimed chimeric genes.

Appellants highlight a quote from the preceding passage: "The 'written description' requirement must be applied in the context of the particular invention and the state of the knowledge. The Board's rule that the nucleotide sequences of the chimeric genes must be fully presented, although the nucleotide sequences of the component DNA are known, is an inappropriate

generalization. When the prior art includes the nucleotide information, precedent does not set a *per se* rule that the information must be determined afresh." Here, the examiner is making a similar requirement – that applicants provide a full description of sequences that are known and not the central feature of the invention. And just as the Federal Circuit found with respect to the Board decision in the *Capron* case, that requirement is improper.

Combinations. It is not clear that the examiner has maintained that aspect of the rejection regarding combinations. As explained previously, the combinations are not ***joined*** peptides but, rather, a set of distinct peptides that may be combined, for example, in an array. Thus, it is again believed that those of skill in the art readily understand that the recited peptides can be used in combination to identify a plurality of antibodies, each of which binding to distinct peptides.

In light of the preceding arguments and explanation, reversal of the rejection is respectfully requested.

C. Rejection Under 35 U.S.C. §102(b)

Claims 27 and 29 stand rejected over Chen *et al.* to the extent that the reference is said to disclose a peptide comprising SEQ ID NOS: 28-30. Careful review of FIG. (1A) of the reference, cited by the examiner as the relevant disclosure, reveals that the illustrated sequence is not a discrete peptide, but rather, is an illustrated portion of a larger protein – EBNA-I. Thus, the reference fails to satisfy the recitation in claim 27 of a peptide of about 40 amino acids or less.

The examiner now argues simply that "[t]he sequence of Fig. 1A is a protein in which comprises SEQ ID NOS 28, 29, and 30 of the instant claims. Claims 27 and 29 are drawn to a peptide composition comprising a peptide moelcule. Therefore, the peptide sequence with a

larger sequence of Chen meets the limitation of the instant claims." Nothing could be more incorrect.

As repeatedly noted by Appellants, the claims are drawn to "A peptide composition comprising a peptide molecule ***consisting of about forty amino acids or less***" and containing one of the recited peptides. Thus, there is an upper limit on the claimed peptides of about 40 amino acids or less. The examiner has admitted that Chen discloses a ***protein***, not a peptide, and that protein is 641 amino acids in length (see Fig. 1A). Thus, Fig. 1A of Chen does not disclose peptides of about forty amino acids or less, and thus cannot anticipate the present claims.

Reversal of the rejection, based on the preceding, is respectfully requested.

IX. Conclusion

In light of the foregoing, appellants respectfully submit that all claims are adequately described and non-obvious over the cited art. Therefore, reversal of all rejections is respectfully requested.

Respectfully submitted,

Date: June 30, 2006

Steven L. Highlander
Reg. No. 37,642
Attorney for Appellants

APPENDIX A – CLAIMS

1-26. (Canceled)

27. (Previously presented) A peptide composition comprising a peptide molecule consisting of about forty amino acids or less and comprising a peptide sequence selected from the group consisting of PPPGRRP (SEQ ID NO:1), GRGRGRGG (SEQ ID NO:2), RGRGREK (SEQ ID NO:3), GPQRRGGDNHGRGRGRGRGGGRPG (SEQ ID NO:13), GGSGSGPRHRDGVRRPQKRP (SEQ ID NO:14), GTGAGAGARGRGG (SEQ ID NO:17), SGGRGRGG (SEQ ID NO:18), RGGSGGRRGRGR (SEQ ID NO:19), SSSSGSPPRRPPPGR (SEQ ID NO:21), RPPPGRRPFFHPVGEADYFEYHQEG (SEQ ID NO:22), GPSTGRPG (SEQ ID NO:25), GQGDGGRRK (SEQ ID NO:26), DGGRRKKGGWFGKHR (SEQ ID NO:27), GKHRGQGGSN (SEQ ID NO:28), GQGGSNPK (SEQ ID NO:29), NPKFENIA (SEQ ID NO:30), RSHVERTT (SEQ ID NO:31), VFVYGGSKT (SEQ ID NO:32), GSKTSLYNL (SEQ ID NO:33), CNIRVTVC (SEQ ID NO:36), PPWFPPMVEG (SEQ ID NO:38) and combinations thereof, wherein the peptide is present either in free form or bound to a carrier molecule.

28. (Canceled)

29. (Previously presented) The composition of claim 27, wherein the peptide molecule or molecules are in a pharmaceutically acceptable carrier.

30. (Withdrawn) The composition of claim 27 mobilized to a solid support.

31. (Withdrawn) The composition of claim 27 labeled with a detectable label.

32. (Withdrawn) The composition of claim 30 immobilized to multiwell plates.

33. (Withdrawn) The composition of claim 30 immobilized to a gel suitable for affinity chromatography.

34. (Withdrawn) The composition of claim 27 bound by autoantibodies in patients characterized by specific disorders.

35. (Withdrawn) A method for determining the likelihood that an individual has or will develop an autoimmune disorder comprising screening their antibodies for reactivity with a peptide molecule consisting of about forty amino acids or less and comprising a peptide sequence selected from the group consisting of PPPGRRP (SEQ ID NO:1), GRGRGRGG (SEQ ID NO:2), RGRGREK (SEQ ID NO:3), GAGAGAGAGAGAGAGAGAGAGAGA (SEQ ID NO:7), GPQRRGGDNHGRGRGRGRGGGRPG (SEQ ID NO:13), GGSGSGPRHRDGVRRPQKRP (SEQ ID NO:14), RPQKRPSC (SEQ ID NO:15), QKRPSCIGCKGTHGGTG (SEQ ID NO:16), GTGAGAGARGRGG (SEQ ID NO:17), SGGRGRGG (SEQ ID NO:18), RGGSGGRRGRGR (SEQ ID NO:19), RARGRGRGRGEKRPRS (SEQ ID NO:20), SSSSGSPPRRPPPGR (SEQ ID NO:21), RPPPGRRPFFHPVGEADYFEYHQEG (SEQ ID NO:22), GPSTGRPG (SEQ ID NO:25), GQGDGGRRK (SEQ ID NO:26), DGGRRKKGGWFGKHR (SEQ ID NO:27), GKHRGQGGSN (SEQ ID NO:28), GQGGSNPK (SEQ ID NO:29), NPKFENIA (SEQ ID NO:30), RSHVERTT (SEQ ID NO:31), VFVYGGSKT (SEQ ID NO:32), GSKTSLYNL (SEQ ID NO:33), GMAPGPGP (SEQ ID NO:34), PQPGPLRE (SEQ ID NO:35), CNIRVTVC (SEQ ID NO:36), RVTVCSFDDG (SEQ ID NO:37), PPWFPPMVEG (SEQ ID NO:38) and combinations thereof, wherein the peptide is present in either free form or bound to a carrier molecule.

36. (Withdrawn) The method of claim 35 wherein the peptide molecules are immobilized to a solid support.

37. (Withdrawn) The method of claim 35 wherein the peptide molecules are labeled with a detectable label.

38. (Withdrawn) The method of claim 35 wherein the peptide molecules are immoblized to multiwell plates.

39. (Withdrawn) The method of claim 35 wherein the peptide molecules are immobilized to a gel suitable for affinity chromatography.

40. (Withdrawn) The method of claim 35 wherein the peptide molecules are bound by autoantibodies.

APPENDIX B – EVIDENCE

Exhibit 1 – Chen *et al.*, *Virology* 205:486-495 (1994)

VIROLOGY 205, 486–495 (1994)

Delineation of a 16 Amino Acid Sequence That Forms a Core DNA Recognition Motif in the Epstein–Barr Virus EBNA-1 Protein

MEI-RU CHEN,* JIANCHAO ZONG,* AND S. DIANE HAYWARD*,†,[1]

*Molecular Virology Laboratories, *Departments of Oncology and †Pharmacology and Molecular Sciences, Johns Hopkins School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205-2185*

Received June 21, 1994; accepted September 7, 1994

EBNA-1 is the sole virally specified protein required for replication of the Epstein–Barr virus latency origin and also modulates its own expression. Both of these functions are mediated through specific DNA binding. We created a series of amino acid substitutions across the previously identified DNA-binding domain of EBNA-1 to further define the amino acids required for DNA recognition. The results of electrophoretic mobility shift assays using *in vitro*-translated EBNA-1 (aa 408–641) polypeptides indicated that: (i) The predicted alpha helical segment between aa 477 and 487 is not directly involved in DNA recognition but appears to contribute to a critical local polypeptide conformation. (ii) The positively charged residues Arg459, Lys460, and Lys461 are dispensable for DNA binding. (iii) The region between Gly462 and Lys477 contains residues important for DNA recognition. This region is extremely sensitive to mutation. The behavior of a synthetic peptide representing EBNA-1 aa 458–478 supported the mutagenesis data. A dimer form of this peptide, which lacks the predicted alpha helical domain, was capable of binding DNA, but only nonspecifically. The effect of binding affinity on transactivation was examined in cotransfection assays. EBNA-1 mutants with reduced binding affinity also demonstrated reduced levels of transactivation.

INTRODUCTION

Replication of the latent, episomal form of the Epstein–Barr virus (EBV) genome requires only a single EBV-encoded protein, EBNA-1 (Yates *et al.*, 1984). The latency origin of replication, oriP, contains multiple EBNA-1-binding sites. OriP can be divided into two separate functional regions (Reisman *et al.*, 1985). The family of repeats, which consists of 20 tandem copies of a 30-bp repeat unit each containing a single EBNA-1-binding site (Jones *et al.*, 1989; Rawlins *et al.*, 1985), acts as an EBNA-1 dependent transcriptional enhancer when placed upstream or downstream of heterologous promoters (Reisman and Sugden, 1986) and stabilizes retention of linked DNA sequences (Krysan *et al.*, 1989; Middleton and Sugden, 1994). At least 7 copies of the EBNA-1-binding sites are needed for the enhancer to function effectively in transactivation, replication, and plasmid maintenance (Wysokenski and Yates, 1989). The enhancer region is required for oriP replication in most settings but replication in the absence of these sequences has been detected in transient replication assays performed in D98/Raji and HeLa cells (Harrison *et al.*, 1994). The second functional region, the dyad symmetry, contains two pairs of lower affinity EBNA-1-binding sites and is the region in which replication is initiated (Gahn and Schildkraut, 1989).

EBNA-1 is not a helicase, nor does it possess any other detectable enzymatic activity (Frappier and O'Donnell, 1992; Middleton and Sugden, 1992; Shah *et al.*, 1992). Binding of EBNA-1 to its sites in the family of repeats region and the dyad symmetry leads to a looping out of the intervening DNA sequences that can be visualized by electron microscopy (Frappier and O'Donnell, 1992; Su *et al.*, 1991). Further, binding of EBNA-1 independently to either region generates significant nuclease hypersensitivity between binding sites, suggesting that EBNA-1 may also influence DNA structure locally (Rawlins *et al.*, 1985). Studies using potassium permanganate footprinting have shown a specific EBNA-1-induced DNA distortion at two thymidines within the dyad symmetry region and this distortion occurs when EBNA-1 is bound both *in vitro* (Frappier and O'Donnell, 1992; Hearing *et al.*, 1992) and *in vivo* (Hsieh *et al.*, 1993). The role of EBNA-1 in oriP replication thus appears to be primarily in the creation of an active origin structure with facilitation of replication complex assembly through protein:protein contacts also a possibility. In addition to its replication function, EBNA-1 also contributes to the regulation of its own expression. EBNA-1 transactivation of expression from the latency C promoter has been shown to occur in the presence of the upstream oriP binding sites (Sugden and Warren, 1989), and immediately downstream of the latency F promoter (Sample *et al.*, 1991; Schaefer *et al.*, 1991; Smith and Griffin, 1992) are two low affinity EBNA-1-binding sites (the Q locus) which may negatively modulate the expression of EBNA-1 from this alternative

[1] To whom correspondence and reprint requests should be addressed. Fax: (410) 955-8685.

0042-6822/94 $6.00

latency promoter (Ambinder *et al.*, 1990; Jones *et al.*, 1989; Sample *et al.*, 1992).

Since DNA binding is essential for EBNA-1's replication and gene regulation functions (Yates and Camiolo, 1988; Polvino-Bodnar and Schaffer, 1992), efforts have been made to characterize this interaction. The EBNA-1 binding sites in the three loci in the EBV genome differ in their affinity for EBNA-1 with the sites in the family of repeats having the highest affinity and those in the Q locus the lowest (Jones *et al.*, 1989). A detailed analysis of the EBNA-1 binding site determined that these differences resulted from sequence variations within the different sites (Ambinder *et al.*, 1990). The EBNA-1 binding site is palindromic and alteration of the spacing between the two half sites abolishes binding (Ambinder *et al.*, 1990). In light of the nature of the binding site it was not surprising when EBNA-1 was subsequently shown to dimerize (Ambinder *et al.*, 1991) and dimerization was found to be essential for binding (Chen *et al.*, 1993). Both DNA-binding and dimerization functions were originally mapped to a segment of EBNA-1 between aa 459 and 607 (Ambinder *et al.*, 1991). Separation of this region into functional subdomains proved to be challenging and was only recently achieved. A deletional analysis across this region identified a bipartite dimerization domain located between aa 501–532 and aa 554–598 and implicated an upstream region between aa 459 and 487 in DNA binding (Chen *et al.*, 1993). The overall organization of the DNA-binding and dimerization domains does not match that of other described DNA-binding proteins. We have now performed a mutational analysis across amino acids 459 to 487 in order to better understand the features of this region that are important for specific DNA binding. The generation of mutated EBNA-1 proteins with reduced DNA-binding affinities has also enabled an examination of the effects of binding affinity on one of EBNA-1's functions, namely transactivation.

MATERIALS AND METHODS

Plasmid constructions

Wild-type EBNA-1 (aa 408 to 641) (referred to hereafter as EBNA-1(408–641)) was expressed by *in vitro* transcription and translation from pRA362 (1). EBNA-1 (459 to 641) was generated similarly from pWS61 (Shah *et al.*, 1992). Plasmids pMRC24 (485D), pMRC26 (477A-SR), and pMRC49 (Δ489-93) have been described previously (Chen *et al.*, 1993). The primer, CTAGAGATCTGGAGGGTGGTTTGGA, and the reverse primer, LGH312 (5′-GGCCCTGCAGTCACTCCTGCCCTTCCTC), were used to amplify aa 462 to 641 of EBNA-1. The purified polymerase chain reaction (PCR) product was then cloned into the *in vitro* transcription vector pGH254 to make pMRC77 (EBNA-1(462–641)). The primers 5′-CTAGAGATCTTGGTTTGGAAAGCAT and LGH 312 were used to amplify aa 464 to 641 of EBNA-1 to generate pMRC103 (EBNA-1(464–641)). The primers 5′-AGGAGAGCTCCTAAACCTTCTGCAA and LGH312 were used to amplify aa 485 to 641 with a mutation at 486, and the purified PCR product was used to replace the same region of pMRC28 to generate pMRC94 (486G).

The first set of EBNA-1 mutants was generated using PCR primers that were synthesized with four possible base substitutions at selected positions. These primers and reverse primer LGH312 were used to amplify aa 481 to 641 of EBNA-1 to generate different mutations between aa 481 to 483. The purified PCR products were cloned into *Xba*I/*Pst*I-cleaved pMRC32 (which contains the mutation 479SR). Individual clones were sequenced and contained the following mutations: pMRC78 (479SRS-A), pMRC79 (479SRSP), pMRC80 (479SRFE), pMRC81 (479SRSPP), pMRC82 (479SRF), pMRC84 (479SR--P), pMRC85 (479SRFQA), pMRC86 (479SRS-Q), and pMRC87 (479SR--A). To generate mutations between aa 476 and 478, a similar approach was taken using primers LRA27 (5′-CCGGATCCTGTAGGGGAAGCCGAT), and a primer with randomized substitutions to amplify aa 408 to 479 of EBNA-1. The PCR products were cloned into the *Xba*I site of pMRC32. After sequencing, the mutations selected were pMRC88 (478ISR), pMRC89 (478NSR), and pMRC90 (476A-YSR).

Three EBNA-1 mutants were generated in the background of pMRC35 (467VD) (Chen, 1993). Primer 5′-TAATACGACTCACTATAGGG, which anneals to T7 promoter sequence, was used with reverse primer 5′-CTAGGTCGACTACAGACCACCCTCCTTTTTT or 5′-CTAGGTCGACTCCAGATATCCCTCCTTTTTTGC to amplify sequences between the T7 promoter and aa 468 of EBNA-1. The PCR products containing mutations at aa 465,466 or aa 464,465 were reintroduced into pMRC35 to generate pMRC96 (464VS-VD) and pMRC97 (465SVVD). Primer 5′-CTAGGTCGACCGTACTTATGGAGGTTCCAAGGG and reverse primer LGH312 were used to amplify a DNA fragment between aa 467 and 641 of EBNA-1 and introduce mutations at aa 470 and 471. The DNA was ligated into *Sal*I/*Pst*I-cleaved pMRC35 to generate pMRC99 (467VD-TY).

Another six EBNA-1 mutants were also generated by PCR mutagenesis. The oligonucleotides LRA27 and 5′-GACCTCTAGACTTTCCAAACCACCC were used to amplify codons 408 to 470 with a mutation at codon 468, and oligonucleotides 5′-AAAGTCTAGAGGTCAAGGAGGT and LGH312 were used to amplify codons 467 to 641 with the same mutation at codon 468. The two PCR products were then purified, denatured, annealed, and again amplified with the outside primers LRA27 and LGH312. The isolated DNA fragment was cloned into pGH253 to create pMRC76 (468S). A similar strategy was used to create pMRC98 (474VD), pMRC104 (464VD), pMRC105 (466VD), pMRC106 (469VD), and pMRC107 (473VD). The primer pairs used to introduce the mutations were 5′-AGGTGTCGACCCGAAATTTGAGAAC and

5′-TCGGGTCGACACCTCCTTGACCACG (474VD), 5′-AGGGGTCGACGGAAAGCATCGTGGT and 5′-TTCCGTCGACCCCTCCTTTTTGCG (464VD), 5′-GTTTGTCGACCATCGTGGTCAAGGA and GATGGTCGACAAACCACCCTCCTTT (466VD), 5′-GCATGTCGACCAAGGAGGTTCCAAC and 5′-CTTGGTCGACATGCTTTCCAAACCA (469VD), and 5′-TCAAGTCGACTCCAACCCGAAATTT and 5′-TGGAGTCGACTTGACCACGATGCTT (473VD).

A modified tissue-culture expression vector (pMRC69) was created by introducing a *Pst*I site into pSG5 which contains the SV40 early promoter and β-globin splicing signal. The primers 5′-CTAGAGATCTATGTCTGACGAGGGGCCA and LGH312 were used to amplify EBNA-1 from p367 (which contains a deletion between aa 102 and 325 (Yates and Camiolo, 1988) and was a gift from John Yates, Roswell Park Cancer Institute, Buffalo, NY). *Bst*XI/*Pst*I DNA fragments from pMRC24, pMRC26, and pMRC49 were cloned into pMRC69 to generate pMRC73 (485D), pMRC74 (477A-SR), and pMRC75 (Δ489-97), respectively. The target plasmid pFRTK-CAT used in the cotransfection assays was a gift from B. Sugden (McArdle Cancer Centre, Madison, WI) (Reisman and Sugden, 1986).

In vitro transcription–translation

Plasmid DNA was linearized downstream of the coding sequence of EBNA-1 and capped mRNA was prepared using an *in vitro* transcription kit (Straragene, La Jolla, CA). *In vitro* translation was performed using rabbit reticulocyte lysate (Promega, Madison, WI) A standard *in vitro* translation reaction contained 1 to 2 μg of mRNA in a 50-μl reaction mixture and 50 μCi of [^{35}S]methionine (800 Ci/mmol) purchased from New England Nuclear (Wilmington, DE). All polypeptides were examined by sodium dodecyl sulfate (SDS)–PAGE for quantitation and to ensure that the appropriately sized product was generated.

Preparation of oligonucleotide-binding substrates and electrophoretic mobility shift assay

Symmetric consensus and mutant-binding site probes were prepared from 19-mer synthetic single-stranded oligonucleotides which can self anneal and thus serve as primers for a Klenow fill-in reaction, generating double-stranded 30-mers (Ambinder *et al.*, 1990). The oligonucleotides were labeled and purified as described (Chen *et al.*, 1993). Equal amounts of the *in vitro*-translated EBNA-1 polypeptides were incubated with probe DNA (20 fmol) for 30 min at room temperature in a 25-μl total volume of 25 m*M* HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (pH 7.5)), 1 m*M* dithiothreitol, 5 m*M* $MgCl_2$, 100 m*M* KCl, 0.1% Nonidet P-40, 1 μg of poly (dI·dC). The reaction mixtures were loaded onto a 6% polyacrylamide gel in HEE buffer (10 m*M* HEPES (pH 7.5), 1 m*M* EDTA, 5 m*M* EGTA) and electrophoresed at 10 V/cm at room temperature for 1.5 hr. The gels were dried and autoradiographed.

DNA transfection and CAT assays

CV-1 monkey kidney cells were grown in Dulbecco's modified Eagle's minimal essential medium supplemented with 10% fetal calf serum. For chloramphenicol acetyltransferase (CAT) assays, CV-1 cells were plated in six-well cluster dishes at 2×10^5 cells per well the day before transfection and were transfected using calcium phosphate-BES [*N*′,*N*-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid]-buffered saline (Chen and Okayama, 1987). The plasmid pSV2–β-gal was cotransfected as an internal control for transfection efficiency. Cells were harvested 40 hr after transfection and assays for CAT and β-galactosidase were performed as previously described (Ling *et al.*, 1993).

Western immunoblot analysis

Cells were lysed 60 hr after transfection in 2× sample buffer (50 m*M* TRIS, 4% SDS, 20% glycerol, 0.04% bromophenol blue, 200 m*M* dithiothreitol) at 2×10^4 cells per microliter and were sonicated briefly. The lysate was electrophoresed through a 10% polyacrylamide gel. After blocking in TRIS-buffered saline (TBS)–5% nonfat dry milk–0.1% Tween 20 for 1 hr and washing with TBS–0.1% Tween, the filter was then incubated with the EBNA-1 monoclonal antibody EBNA-OT1× (Chen *et al.*, 1993) in TBS–0.1% Tween at room temperature for 1 hr. The filter was washed three times for 15 min in TBS–Tween and then incubated in a 1:5000 dilution of alkaline phosphatase-conjugated goat anti-mouse antibody. The filter was washed three times and developed with an ECL kit (Pharmacia, Inc., Piscataway, NJ) followed by exposure to Kodak X-ray film.

DNA-binding of a synthetic peptide dimer

A peptide corresponding to residues 458 to 478 of EBNA-1 was synthesized with a Gly–Gly–Cys (Talanian *et al.*, 1990) linker added at the carboxyl terminus using an Applied Biosystems Model 430A peptide synthesizer and purified by reverse-phase HPLC. A mixture of 15 mg of peptide monomer and 180 mg of oxidized DTT was dissolved in 10 ml of 0.1 *M* TRIS·HCl (pH 8.5) by gentle stirring for 24 hr at room temperature. The oxidized dimers were then purified by reverse-phase HPLC. DNA-binding reactions contained 20 m*M* TRIS (pH 7.4), 4 m*M* KCl, 2 m*M* $MgCl_2$, 2 m*M* EDTA, 0.1% NP-40, 2000 cpm of ^{32}P-labeled probe and 0.01–0.25 μ*M* of peptide. Mixtures were incubated at 4° for 30 min and resolved by nondenaturing 8% polyacrylamide gel electrophoresis in TE at 4° (Talanian *et al.*, 1990).

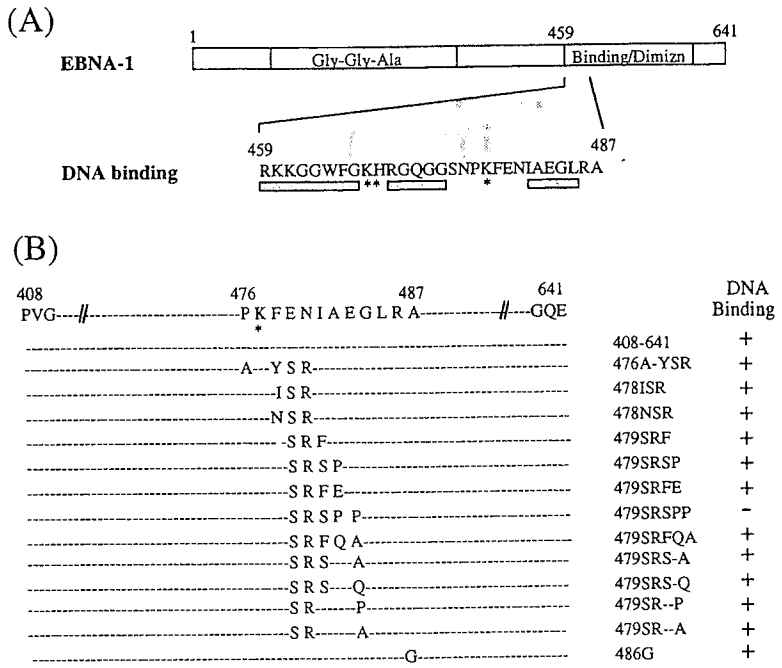

FIG. 1. Organization of the EBNA-1 DNA recognition domain. (A) The location of the DNA recognition domain within the 641 aa EBNA-1 protein is shown along with the amino acid sequence. The locations of deletions known to abolish DNA binding (shaded rectangles) and of mutations known to affect DNA recognition (*) are indicated (Chen *et al.*, 1993). (B) Summary of the structure of 13 newly generated constructions carrying mutations in the predicted alpha helical region (aa 476–487) of the recognition domain. The ability of these constructions to bind to a consensus site is also summarized. The DNA-binding data are shown in Fig. 2A.

RESULTS

The predicted alpha helical segment of the DNA recognition domain is not directly involved in DNA contacts

In a previous study, the DNA-binding domain of EBNA-1 was localized to a region between aa 459 and 487, upstream of the dimerization domain (Chen *et al.*, 1993). This designation was based on the inability of three dimerization competent deletions (Δ459–466, Δ469–473, and Δ481–485) to bind DNA along with the altered pattern of DNA recognition demonstrated by polypeptides carrying substitutions at aa 467, 468, and 477 (summarized in Fig. 1A). The predicted secondary structure of this region (Chou and Fasman, 1974) identified the segment of EBNA-1 between aa 477 and 487 as an alpha helix. A number of characterized DNA-binding proteins have DNA recognition domains which adopt an alpha helical conformation (Harrison, 1991; Harrison and Aggarwal, 1990). We therefore introduced a series of mutations across the aa 477–487 region to assess whether this region of EBNA-1 was directly contributing to site-specific DNA recognition.

We had previously shown that a mutant, EBNA-1(479SR), had DNA-binding properties that were indistinguishable from wild-type protein when assessed against both the consensus EBNA-1-binding site and a series of mutant-binding sites (Chen *et al.*, 1993). The 479SR mutation introduces a unique *Xba*I site. This DNA was used as a template for PCR-directed mutagenesis along with a 5′ primer oligonucleotide that carried randomized substitutions at selected positions. The amplified PCR product thus contained a mixture of mutations. After cloning, individual DNAs were sequenced to identify the particular mutation present. This protocol generated mutations between aa 476 and 483 (476A-YSR to 479SR--A in Fig. 1B). A directed mutation was also introduced at position 486 (486G). The mutant polypeptides, all in an aa 408–641 background, were synthesized by *in vitro* translation and tested for the ability to bind to a 30-mer oligonucleotide consensus-binding site in an electrophoretic mobility shift assay (EMSA). Twelve of the 13 mutants were capable of binding to the consensus binding site (Figs. 1B and 2A). Among those binding was mutant 479SRFQA which contains five consecutive altered amino acids (479–483). Polypeptides carrying mutations at the amino (476A-YSR) and carboxy (486G) boundaries of the region of interest were also able to bind the oligonucleotide probe. We next examined the pattern of binding exhibited by these mutant polypeptides to five previously characterized (Ambinder *et al.*, 1990), lower affinity oligonucleotide probes carrying substitutions at different positions within the binding site. An ability to discriminate between these probes would be indicative of a direct involvement in DNA recognition. No such discrimination was observed. Instead the mutants exhibited one of two patterns. (i) They bound similarly to wild-type EBNA-1(408–641) on all probes tested (five examples)

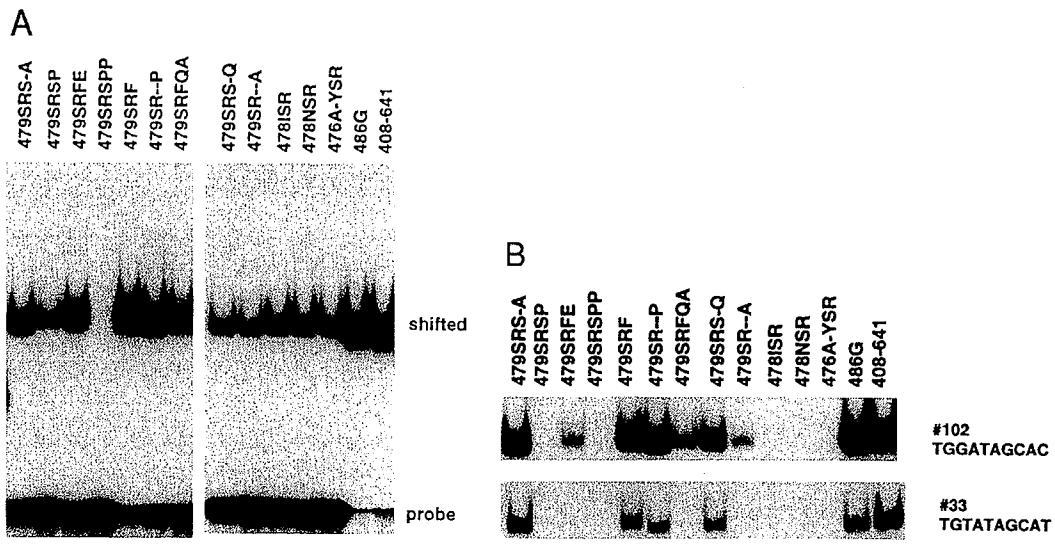

FIG. 2. Effect of mutations in the predicted alpha helical region (aa 476–487) on DNA binding. *In vitro*-translated EBNA-1 polypeptides were incubated with ^{32}P-labeled 30-mer oligonucleotide probes representing (A) high affinity, (B) intermediate affinity (#102) and low affinity (#33) binding sites. Complex formation was analyzed by electrophoresis through 6% polyacrylamide gels. Only the shifted complexes are shown in (B).

or (ii) they demonstrated an exaggerated reduction in binding to the mutated sites that was independent of the position of the substituted base. Four examples bound only to the consensus site while three examples bound to the high/intermediate affinity oligo 102 but not to any of the lower affinity binding sites. The results obtained are illustrated by the EMSA in Fig. 2B which shows the binding pattern to oligonucleotide 102 and the lower affinity oligonucleotide 33. Three of the four mutants that bound only to the consensus-binding site contained a mutation of phenylalanine 478, suggesting that although this amino acid does not appear to directly contact DNA, it may be important in creating an appropriate local polypeptide conformation.

The only mutation to eliminate binding to the consensus site was 479SRSPP. Each of the changes in this mutant are acceptable separately. For example, 479SRSP which contains the same first four altered amino acids is DNA binding and a mutant, 479SR--P, which has the proline substitution at position 483 is also DNA binding. It seems that the combination of Pro–Pro at 482, 483 has an effect that is not generated by the individual presence of proline residues at positions 482 and 483. One possibility is that the introduction of the double proline perturbs the local polypeptide conformation. In particular, two adjacent proline residues would almost certainly disrupt an alpha helical structure whereas a single proline may not be capable of doing so.

The positively charged -RKK- residues at aa 459–461 are dispensable

Previous work (Chen *et al.*, 1993) had implicated Lys467, His468, and Lys477 as residues that may be directly involved in DNA contacts. Consequently, we were interested in the requirement for the positively charged residues Arg459, Lys460, and Lys461. A deletion was constructed that expressed only aa 462–641 and therefore lacked these three residues. Binding of EBNA-1 (462–641) to the consensus binding site and to the low affinity site, oligonucleotide 33, was examined in an EMSA. The 462–641 polypeptide bound the consensus-binding site (Fig. 3A) but not oligonucleotide 33 (Fig. 3B) or any of the other low affinity sites tested. Thus an EBNA-1 polypeptide lacking the Arg–Lys–Lys residues at 459–461 remained capable of DNA binding but the

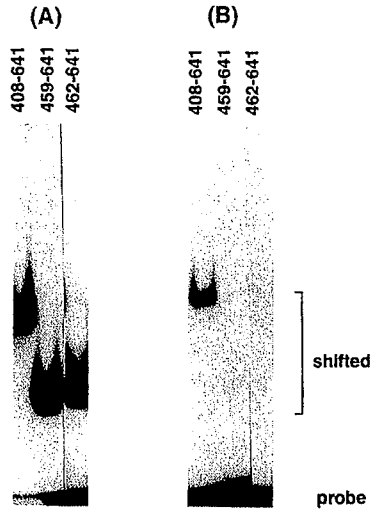

FIG. 3. The positively charged amino acids R459, K460, and K461 are not essential for DNA binding. Electrophoretic mobility shift assays comparing the binding ability of an EBNA-1 polypeptide, 462-641, lacking these three positively charged amino acids with larger polypeptides commencing at aa 459 and aa 408. (A) Binding to the consensus site. (B) Binding to a low affinity site, oligonucleotide #33.

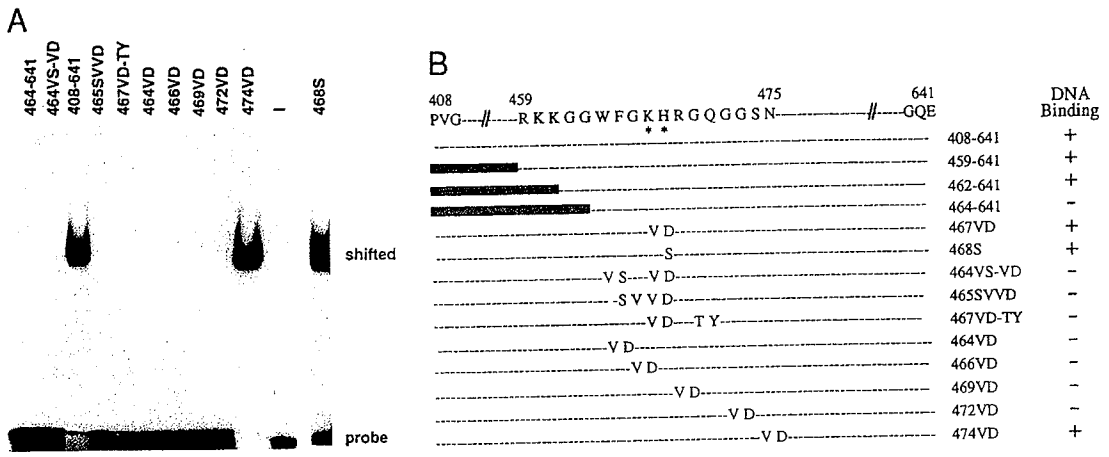

FIG. 4. The region between aa 462 and aa 473 is exquisitely sensitive to mutation. (A) Electrophoretic mobility shift assay assessing the effect of mutation in this region on binding to a consensus site probe. (B) Structure and DNA-binding ability of EBNA-1 polypeptides carrying mutations between aa 459 and 475. The data summarized are from Fig. 3, Fig. 4A, and Chen *et al.* (1993; 467VD).

stability of the interaction was reduced compared to EBNA-1(408–641) as evidenced by the inability to bind to low affinity sites. However, EBNA-1(459–641) also bound much more weakly to low affinity sites than EBNA-1(408–641) (Fig. 3), indicating that some of the destabilization might be a generalized effect of placing the amino terminus of the polypeptide very close to the DNA recognition residues.

Mutations between aa 462 and 473 define a locus that is critical for DNA recognition

Having eliminated aa 459–461 and aa 478–487 as being directly involved in DNA recognition, we next examined the requirement for amino acids in the remaining segment of the previously defined DNA-binding domain. Seven of eight paired substitutions introduced into the aa 462–475 region destroyed the ability of the EBNA-1 polypeptides to bind to the consensus-binding site (Fig. 4A).

We had previously shown that a mutant polypeptide 467VD bound the consensus binding site as well as wild-type EBNA-1(408–641) (Chen *et al.*, 1993) and the first three substitution mutations were created in the background of 467VD. In these constructions, the substitution of Val–Ser for Trp464–Phe465; Ser–Val for Phe465–Gly466, or Thr–Tyr for Gln470–Gly471 eliminated DNA binding. Each of these polypeptides carried substitutions at four positions because of the 467VD background. To verify that the 467VD alteration was not influencing the results obtained, we also generated paired mutations across the same region in a wild-type background. In this background, substitution of Val–Asp for Trp464–Phe465, for Gly466–Lys467, for Gly469–Gln470, or for Gly472–Gly473 abolished DNA-binding ability. The only mutant in this set that retained DNA-binding ability was 474VD which substituted valine and aspartate for Ser474 and Asn475.

Deletion of Gly462 and Gly463 in the construction EBNA-1(464–641) also resulted in a nonbinding polypeptide. The lack of DNA-binding ability contrasts with the binding positive EBNA-1(462–641) polypeptide (Fig. 3) which differs only in the addition of the amino-terminal two glycine residues. The data from Figs. 3 and 4A are summarized in Fig. 4B. Overall, the region between aa 462 and aa 473 is extremely intolerant of amino acid substitution. The only substitutions that were compatible with DNA binding were the previously described mutation of aa 467,468 (467VD (Chen *et al.*, 1993)) and a single substitution at aa 468 (468S; Fig. 4A). The mutational data strongly suggests that the aa 462–473 segment of EBNA-1 constitutes part of a DNA recognition domain.

A peptide representing EBNA-1 aa 458–478 can bind DNA as a dimer

The behavior of the EBNA-1 polypeptides carrying amino acid substitutions between aa 462 and aa 473, along with previous data (Chen *et al.*, 1993) on the behavior of a mutation at Lys477, suggested that the amino acids involved in DNA recognition lay between Gly462 and Lys477. We therefore synthesized a polypeptide representing residues 458–478 of EBNA-1 and asked whether this peptide was capable of binding to DNA. Since EBNA-1 binds only in the dimer form (Ambinder *et al.*, 1991; Chen *et al.*, 1993) it was necessary to generate a dimer form of the test peptide. Three residues, Gly–Gly–Cys, were added to the EBNA-1 peptide sequence at the carboxy terminus to provide a spacer plus dimerization capability through the formation of a cystine bridge (Talanian *et al.*, 1990).

As shown in Figs. 5A–5D the HPLC-purified monomer peptide could be converted to the dimer form by incubation with oxidized dithiothreitol. The HPLC-purified dimer could be reconverted to the monomer form by incubation with 10 m*M* dithiothreitol. An EMSA performed with the

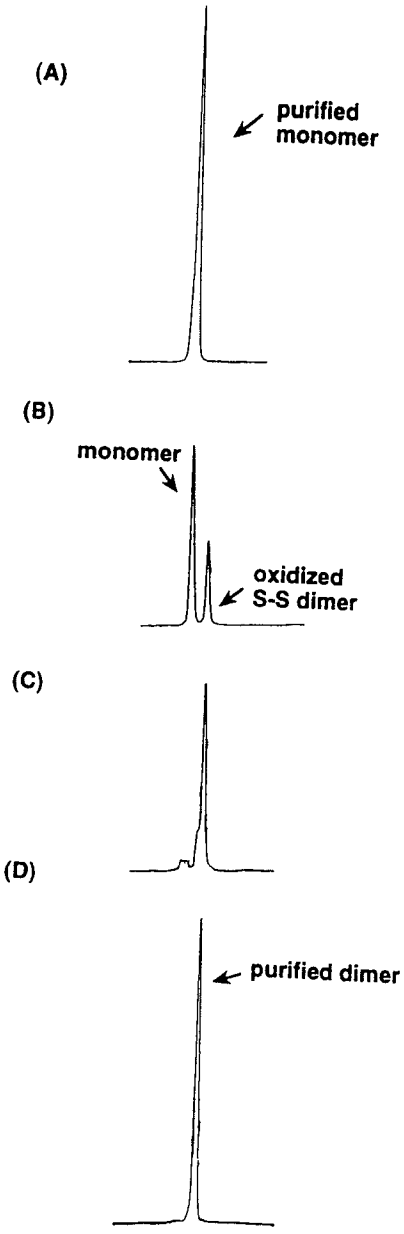

FIG. 5. A peptide representing aa 458–478 and containing the minimal DNA recognition locus can be converted to the dimer form by oxidation of added carboxyterminal cysteine residues. (A–D) Generation and chromatographic separation of monomer and dimer forms of the peptide.

EBNA-1 consensus site as probe indicated that the purified dimer peptide could bind DNA at a peptide concentration of 0.25 μM (Fig. 6). Conversion of the peptide to the monomer form by the addition of 10 mM dithiothreitol abolished the ability of the peptide to bind the probe but did not affect the binding of the *in vitro*-translated EBNA-1(408–641) polypeptide which contains the natural dimerization domain. Although the peptide bound DNA it did not retain the specificity of EBNA-1(408–641). A mutated 30-mer probe which was not bound by EBNA-1 was bound as well as the consensus site by the dimer peptide

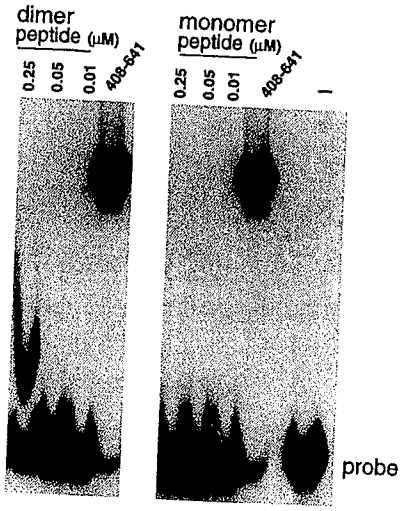

FIG. 6. The dimer aa 458–478 peptide binds DNA. Electrophoretic mobility shift assay demonstrating binding of the dimer peptide at 0.25 μM to a consensus 30-mer site. The monomer form of the peptide did not bind.

at 0.25 μm (data not shown). Thus the region of EBNA-1 represented by the peptide, aa 458–478, has an inherent ability to interact with DNA when in a dimer form but specific site recognition requires additional polypeptide sequences. These would most likely function to place the DNA recognition region in the appropriate conformation to form specific, stable DNA contacts. Our interpretation of the functional organization of the EBNA-1 DNA recognition domain is presented in Fig. 7.

EBNA-1-binding affinity correlates with transactivation ability

There are three loci of EBNA-1-binding sites in the EBV genome: the family of repeats and dyad symmetry loci in oriP and the Q sites located immediately downstream of the latency F promoter (Ambinder *et al.*, 1990; Rawlins *et al.*, 1985). The binding sites in each of these loci differ in their affinity for EBNA-1. The sites in the family of repeats have the highest affinity followed by those in the dyad symmetry and the Q locus has the lowest affinity (Jones *et al.*, 1989). These differences in affinity may serve to modulate EBNA-1 function. The iden-

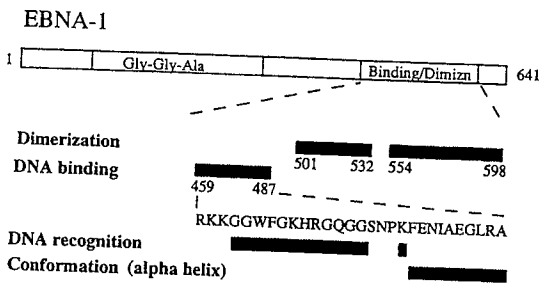

FIG. 7. Summary interpretation of the functional organization of the EBNA-1 DNA-binding domain.

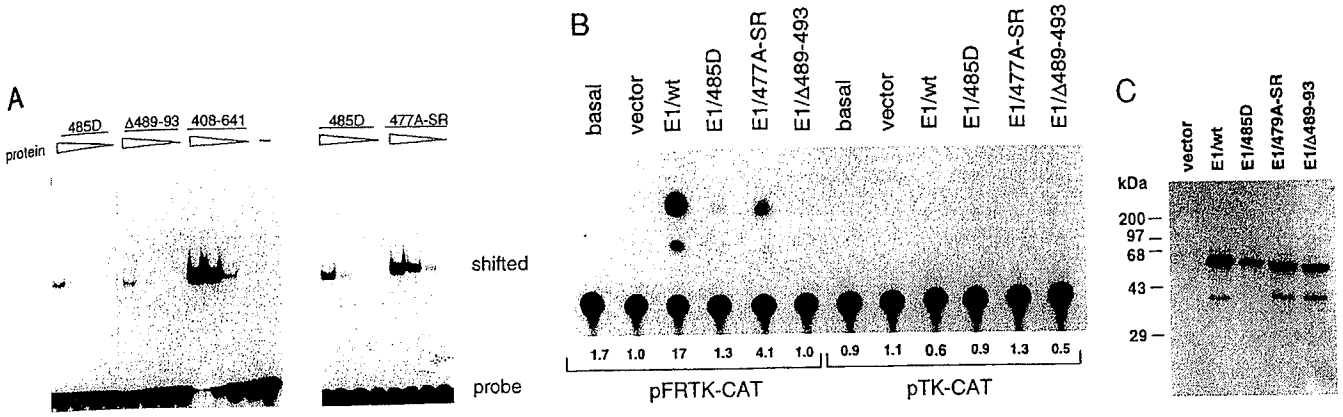

FIG. 8. The transactivation activity of EBNA-1 mutants correlates with their DNA-binding affinity. (A) Electrophoretic mobility shift assays comparing binding of *in vitro*-translated wild-type and mutant EBNA-1 polypeptides to a 30-mer consensus site. Equal amounts of each polypeptide were tested undiluted and at 1:10, 1:100, and 1:1000 dilutions. (B) Cotransfection assay comparing the ability of wild-type EBNA-1 and EBNA-1 proteins carrying the 477A-SR, 485D, and Δ489-93 mutations to activate expression from a target plasmid carrying EBNA-1-binding sites (pFRTK-CAT). The control plasmid pTK-CAT lacks EBNA-1-binding sites. The percentage conversion of [^{14}C]chloramphenicol to the acetylated form is shown beneath each lane. (C) Immunoblot analysis demonstrating that the expression of the different EBNA-1 constructions in transfected cells is comparable.

tification of EBNA-1 mutants with altered affinity for a consensus-binding site (Chen *et al.*, 1993) provides an opportunity to further examine the biological consequences of differences in the affinity of the EBNA-1 interaction with its recognition sequence.

EBNA-1(408–641) and three EBNA-1(408–641) variants carrying mutations that modified, but did not prevent, binding to the consensus binding site from the family of repeats were compared in an EMSA. Equal amounts of each polypeptide were tested at 1, 1:10, and 1:100 dilutions to determine relative binding affinities (Fig. 8A). The EMSA indicated a hierarchy of affinity of EBNA-1(408–641) > 477A-SR > 485D ≥ Δ489-93. The three mutations were incorporated into a eukaryotic EBNA-1 expression vector, p367 (Yates and Camiolo, 1988), that is deleted for the Gly–Gly–Ala repeat region. The transactivation ability of the three EBNA-1 mutants was then compared with that of wild-type EBNA-1 (Δ Gly–Gly–Ala) in a cotransfection assay in Vero cells (Fig. 8B). Cotransfection of the EBNA-1 expression vectors did not significantly affect expression from a control pTK-CAT reporter plasmid. The reporter pFRTK-CAT contains the family of repeats located downstream of the CAT open reading frame (Reisman and Sugden, 1986). Cotransfection of this reporter plasmid with wild-type EBNA-1 (E1/WT) resulted in an approximately 17-fold activation of expression from pFRTK-CAT. The mutant E1/477A-SR produced a fourfold activation and neither of the other mutants, E1/485D or E1/Δ489–493, had a detectable effect. The difference in transactivation was not the result of different levels of expression of the EBNA-1 proteins since their expression in transfected cells was comparable when examined by immunoblotting (Fig. 8C). Thus, a reduced affinity for the consensus binding site, as measured *in vitro* by an EMSA, correlated with a reduction in transactivation ability in transfected tissue culture cells.

DISCUSSION

We had previously implicated a region of EBNA-1 between amino acids 459 and 487 in DNA binding (Chen *et al.*, 1993). This designation was based on the inability of polypeptides carrying deletions between R459 and G466, between R469 and G473 and between I481 and L485 to bind DNA although they retained the ability to dimerize. The arrangement of the EBNA-1 DNA binding and dimerization domains does not conform to that of any recognized class of DNA-binding proteins but an analysis of the predicted secondary structure of the DNA-binding region indicated that the segment between P476 and A487 has a high probability of adopting an alpha helical structure while the remainder of this region appears unstructured. Physical analyses of DNA-binding proteins have identified interaction with DNA through alpha helical domains as a common theme (Harrison and Aggarwal, 1990; Harrison, 1991; Clark *et al.*, 1993; Ellenberger *et al.*, 1992; Hegde *et al.*, 1992; Lee *et al.*, 1993; Omichinski *et al.*, 1993). The primary goal of the current mutagenesis approach was to determine the contribution of the predicted alpha helical region to DNA binding by EBNA-1.

Thirteen mutant EBNA-1 polypeptides were created that carried alterations between amino acids P476 and R486 in the predicted alpha helical region. Despite the fact that multiple alterations were incorporated into 12 of these polypeptides only one failed to bind to a consensus binding site. In the nonbinding mutant A482,E483 were replaced with two adjacent proline residues, a change that would almost certainly disrupt alpha helical structure. The interpretation of structural perturbation as the explanation for the loss of DNA-binding activity is strengthened by the observation that individual proline residues at positions 482 or 483 were not disruptive and

hence the proline residues per se are not incompatible with DNA binding (479SRSP and 479SR--P both bind DNA). A previously described mutation of leucine 485 to a charged aspartate residue also yielded a DNA-binding polypeptide (Chen *et al.*, 1993).

In contrast to the limited effect of mutation on the alpha helical region, mutation of the remainder of the previously defined DNA recognition region, R459 to N475, was highly disruptive. Only 3 of the 11 polypeptides carrying mutations in this region retained DNA-binding activity as assayed by EMSA. Of these three, 2 were at the outer boundaries of the domain (ΔR459,K460,K461 [464–641], and S474,N475 [474VD]) and the third was a single mutation at a position, H468, that had been previously shown to retain DNA-binding ability as part of a paired mutation (467VD; Chen *et al.*, 1993). These results reinforce the conclusions made in the earlier study which incorporated small deletions into the R459 to G473 region and provide more precise definition of the amino acids important for DNA recognition as being those located between G462 and G473 plus the previously identified K477 (Chen *et al.*, 1993).

Positively charged amino acids are frequently important components of DNA recognition domains making contacts either with specific bases or with the phosphate backbone. The RKK grouping at aa 459–461 is dispensable for DNA binding but mutation of K467, R469, or K477 either prevents binding or alters binding specificity, suggesting a role for these residues in specific recognition. A striking feature of the G462 to K477 DNA recognition domain is the presence of six glycines. Glycine residues introduce conformational flexibility. EBNA-1 is known to change conformation on binding DNA (Shah *et al.*, 1992). It seems likely that the appropriate structure for specific DNA interaction is imposed by the adjacent downstream residues. Upstream residues are apparently less important since a polypeptide commencing at G462 is capable of DNA binding.

A synthetic peptide representing aa 458–478 that contains the DNA recognition region but lacks the predicted alpha helical segment was able to bind DNA in the dimer form but did not retain EBNA-1-binding specificity. The addition of cysteine residues to synthetic peptides to generate a dimer DNA-binding species was used successfully in the study of the DNA recognition domain of the yeast transcriptional activator GCN4 (Talanian *et al.*, 1990). In this case, reconstitution of DNA-binding activity may be possible because the cystine bridge recapitulates the natural hinge between the GCN4 dimerization and DNA recognition domains. The structure of these two domains in bZIP proteins is represented by the scissors grip model (Vinson *et al.*, 1989; Ellenberger *et al.*, 1992). The lack of binding specificity of the EBNA-1 peptide implies that the region between the dimerization and DNA recognition regions of EBNA-1 adopts a very different configuration. The inability of the peptide to bind with specificity also emphasizes the importance of the adjacent helical region. The overall model for EBNA-1:DNA interaction that emerges from these studies is one in which residues between G462 and K477 interact with DNA and in which the adoption of the appropriate structure for specific DNA recognition is strongly influenced by the alpha helical nature of the adjacent F478 to A487 region and also potentially by the downstream dimerization domain (Fig. 7). The proximity of the DNA recognition sequences to the adjacent alpha helical region is reminiscent of the arrangement in basic helix–loop–helix proteins. However, in that family of proteins the adjacent alpha helix forms part of the dimerization domain whereas in EBNA-1 it does not (Chen *et al.*, 1993).

The binding sites in the EBV genome display a hierarchy of binding affinity for EBNA-1 ranging from the high affinity sites in the family of repeats to the low affinity sites at the Q locus (Jones *et al.*, 1989). In a transient expression assay, EBNA-1 mutants with reduced binding affinity were also less effective transactivators. Thus, binding site affinity can influence EBNA-1 function. This result is consistent with the data of Wysokenski and Yates (1989) who showed that EBNA-1 transactivation of a target containing binding sites from the family of repeats was greater than that observed if the target plasmid contained the same number of binding sites from the dyad symmetry region. Further, our results indicate that it is the binding site affinity that dictates the outcome for transactivation and not the relative spacing of the sites which also differs at the different loci.

ACKNOWLEDGMENTS

The EBNA-OT1x monoclonal antibody was kindly provided by Jaap Middledorp, Organon Teknika. We thank Wu-Schyong Liu for assistance with the synthetic peptide purification, Mabel Chiu for technical support, and Feng Chang for manuscript preparation. This work was supported by Public Health Service Grants RO1 CA42243 (S.D.H.) and AI31454 (G.S.H.) from the National Institutes of Health and an American Cancer Society Faculty Research Award, FRA-429.

REFERENCES

AMBINDER, R. F., SHAH, W. A., RAWLINS, D. R., HAYWARD, G. S., and HAYWARD, S. D. (1990). Definition of the sequence requirements for binding of the EBNA-1 protein to its palindromic target sites in Epstein–Barr virus DNA. *J. Virol.* **64**, 2369–2379.

AMBINDER, R. F., MULLEN, M., CHANG, Y.-N., HAYWARD, G. S., and HAYWARD, S. D. (1991). Functional domains of Epstein–Barr virus nuclear antigen EBNA-1. *J. Virol.* **65**, 1466–1478.

CHEN, C., and OKAYAMA, H. (1987). High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752.

CHEN, M.-R., MIDDELDORP, J. M., and HAYWARD, S. D. (1993). Separation of the complex DNA binding domain of EBNA-1 into DNA recognition and dimerization subdomains of novel structure. *J. Virol.* **67**, 4875–4885.

CHOU, P. Y., and FASMAN, G. D. (1974). Prediction of protein conformation. *Biochemistry* **13**, 222–245.

CLARK, K. L., HALAY, E. D., LAI, E., and BURLEY, S. K. (1993). Co-crystal structure of the HNF-3/*fork head* DNA-recognition motif resembles histone H5. *Nature* **364**, 412–420.

ELLENBERGER, T. E., BRANDL, C. J., STRUHL, K., and HARRISON, S. C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: Crystal structure of the protein–DNA complex. *Cell* 71, 1223–1237.

FRAPPIER, L., and O'DONNELL, M. (1991). Overproduction, purification, and characterization of EBNA-1, the origin binding protein of EBV. *J. Biol. Chem.* 266, 7819–7826.

FRAPPIER, L., and O'DONNELL, M. (1991). Epstein–Barr virus nuclear antigen 1 mediates a DNA loop within the latent replication origin of Epstein–Barr virus. *Proc. Natl. Acad. Sci. USA* 88, 10875–10879.

FRAPPIER, L., and O'DONNELL, M. (1992). EBNA-1 distorts oriP, the Epstein–Barr virus latent replication origin. *J. Virol.* 66, 1786–1790.

GAHN, T. A., and SCHILDKRAUT, C. L. (1989). The Epstein–Barr virus origin of plasmid replication, ori-P, contains both the initiation and termination sites of DNA replication. *Cell* 58, 527–535.

HARRISON, S. C., and AGGARWAL, A. K. (1990). DNA recognition by proteins with the helix–turn–helix motif. *Annu. Rev. Biochem.* 59, 933–969.

HARRISON, S. C. (1991). A structural taxonomy of DNA-binding domains. *Nature* 353, 715–719.

HARRISON, S., FISENNE, K., and HEARING, J. (1994). Sequence requirements of the Epstein–Barr virus latent origin of DNA replication. *J. Virol.* 68, 1913–1925.

HEARING, J., MULHAUPT, Y., and HARPER, S. (1992). Interaction of Epstein–Barr virus nuclear antigen 1 with the viral latent origin of replication. *J. Virol.* 66, 694–705.

HEGDE, R. S., GROSSMAN, S. R., LAIMINS, L. A., and SIGLER, P. B. (1992). Crystal structure at 1.7 A of the bovine papillomavirus-1 E1 DNA-binding domain bound to its DNA target. *Nature* 359, 505–512.

HSIEH, D.-J., CAMIOLO, S. M., and YATES, J. L. (1993). Constitutive binding of EBNA-1 protein to the EBV replication origin, oriP, with distortion of DNA structure during latent infection. *EMBO. J.* 12, 4933–4944.

JONES, C. H., HAYWARD, S. D., and RAWLINS, D. R. (1989). Interaction of lymphocyte-derived Epstein–Barr virus nuclear antigen EBNA-1 with its DNA binding sites. *J. Virol.* 63, 101–110.

KRYSAN, P. J., HAASE, S. B., and CALOS, M. P. (1989). Isolation of human sequences that replicate autonomously in human cells. *Mol. Cell. Biol.* 9, 1026–1033.

LEE, M. S., KLIEWER, S. A., PROVENCAL, J., WRIGHT, P. E., and EVANS, R. M. (1993). Structure of the retinoid X receptor α DNA binding domain: A helix required for homodimeric DNA binding. *Science* 260, 1117–1121.

LING, P. D., RYON, J. J., and HAYWARD, S. D. (1993). EBNA-2 of herpesvirus papio diverges significantly from the type A and type B EBNA-2 proteins of Epstein–Barr virus but retains an efficient transactivation domain with a conserved hydrophobic motif. *J. Virol.* 67, 2990–3003.

MIDDLETON, T., and SUGDEN, B. (1992). EBNA-1 can link the enhancer element to the initiator element of the Epstein–Barr virus plasmid origin of DNA replication. *J. Virol.* 66, 489–495.

MIDDLETON, T., and SUGDEN, B. (1994). Retention of plasmid DNA in mammalian cells is enhanced by binding of the Epstein–Barr virus replication protein EBNA1. *J. Virol.* 68, 4067–4071.

OMICHINSKI, J. G., CLORE, G. M., SCHAAD, O., FELSENFELD, G., TRAINOR, C., APPELLA, E., STAHL, S. J., and GRONENBORN, A. M. (1993). NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* 261, 438–446.

POLVINO-BODNAR, M., and SCHAFFER, P. A. (1992). DNA binding activity is required for EBNA 1-dependent transcriptional activation and DNA replication. *Virology* 187, 591–603.

RAWLINS, D. R., MILMAN, G., HAYWARD, S. D., and HAYWARD, G. S. (1985). Sequence specific DNA-binding of EBNA-1 to clustered sites in the plasmid maintenance region. *Cell* 42, 859–868.

REISMAN, D., YATES, J., and SUGDEN, B. (1985). A putative origin of replication of plasmids derived from Epstein–Barr virus is composed of two cis-acting components. *Mol. Cell. Biol.* 5, 1822–1832.

REISMAN, D., and SUGDEN, B. (1986). Transactivation of an Epstein–Barr viral transcriptional enhancer by the Epstein–Barr viral nuclear antigen 1. *Mol. Cell. Biol.* 6, 3838–3846.

SAMPLE, J., BROOKS, L., SAMPLE, C., YOUNG, L., ROWE, M., GREGORY, C., RICKINSON, A., and KIEFF, E. (1991). Restricted EBV protein expression in Burkitt lymphoma is due to a different Epstein–Barr nuclear antigen 1 transcriptional initiation site. *Proc. Natl. Acad. Sci. USA* 88, 6343–6347.

SAMPLE, J., HENSON, E. B. D., and SAMPLE, C. (1992). The Epstein–Barr virus nuclear protein 1 promoter active in type 1 latency is autoregulated. *J. Virol.* 66, 4654–4661.

SCHAEFER, B. C., WOISETSCHLAEGER, M., STROMINGER, J. L., and SPECK, S. H. (1991). Exclusive expression of EBV nuclear antigen 1 in Burkitt lymphoma arises from a third promoter, distinct from the promoters used in latently infected lymphocytes. *Proc. Natl. Acad. Sci. USA* 88, 6550–6554.

SHAH, W. A., AMBINDER, R. F., HAYWARD, G. S., and HAYWARD, S. D. (1992). Binding of EBNA-1 to DNA creates a protease-resistant domain that encompasses the DNA recognition and dimerization functions. *J. Virol.* 66, 3355–3362.

SMITH, P. R., and GRIFFIN, B. E. (1992). Transcription of the Epstein–Barr virus gene, EBNA-1 from different promoters in nasopharyngeal carcinoma and B-lymphoblastoid cells. *J. Virol.* 66, 706–714.

SU, W., MIDDLETON, T., SUGDEN, B., and ECHOLS, H. (1991). DNA looping between the origin of replication of Epstein–Barr virus and its enhancer site: Stabilization of an origin complex with Epstein–Barr nuclear antigen 1. *Proc. Natl. Acad. Sci. USA* 88, 10870–10874.

SUGDEN, B., and WARREN, N. (1989). A promoter of Epstein–Barr virus that can function during latent infection can be transactivated by EBNA-1, a viral protein required for viral DNA replication during latent infection. *J. Virol.* 63, 2644–2649.

TALANIAN, R. V., MCKNIGHT, C. J., and KIM, P. S. (1990). Sequence-Specific DNA binding by a short peptide dimer. *Science* 249, 769–770.

VINSON, C. R., SIGLER, P. B., and MCKNIGHT, S. L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science* 246, 911–916.

WILLIAMS, D. L., and KOWALSKI, D. (1993). Easily unwound DNA sequences and hairpin structures in the Epstein–Barr virus origin of plasmid replication. *J. Virol.* 67, 2707–2715.

WYSOKENSKI, D. A., and YATES, J. L. (1989). Multiple EBNA-1 binding sites are required to form an EBNA-1 dependent enhancer and to activate a minimal replicative origin within oriP of Epstein–Barr virus. *J. Virol.* 63, 2657–2666.

YATES, J., WARREN, N., REISMAN, D., and SUGDEN, B. (1984). A cis acting element from the Epstein–Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. *Proc. Natl. Acad. Sci. USA* 81, 3806–3810.

YATES, J. L., WARREN, N., and SUGDEN, B. (1985). Stable replication of plasmids derived from Epstein–Barr virus in a variety of mammalian cells. *Nature* 313, 812–815.

YATES, J. L., and CAMIOLO, S. M. (1988). Dissection of DNA replication and enhancer activation functions of Epstein–Barr virus nuclear antigen 1. *Cancer Cells* 6, 197–205.

APPENDIX C – RELATED PROCEEDINGS

None